

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/576,030
Source: 1Fwp
Date Processed by STIC: 4/27/06

ENTERED



IFWP

RAW SEQUENCE LISTING

DATE: 04/27/2006

PATENT APPLICATION: US/10/576,030

TIME: 07:23:05

Input Set : A:\toyal49.001apc.txt

Output Set : N:\CRF4\04272006\J576030.raw

4 <110> APPLICANT: Yamaguchi, Masanori
 5 Kita, Yuichi
 6 Mori, Tetsuya
 7 Kanbe, Kenji
 8 Tomoda, Akihiro
 9 Takahashi, Atsushi
 10 Ichikawa, Wakako
 12 <120> TITLE OF INVENTION: METHOD FOR PRODUCING SCYLLO-INOSITOL
 14 <130> FILE REFERENCE: TOYA149.001APC
 C--> 16 <140> CURRENT APPLICATION NUMBER: US/10/576,030
 C--> 17 <141> CURRENT FILING DATE: 2006-04-13
 19 <150> PRIOR APPLICATION NUMBER: PCT/JP2004/015174
 20 <151> PRIOR FILING DATE: 2004-10-14
 22 <150> PRIOR APPLICATION NUMBER: JP2003-353490
 23 <151> PRIOR FILING DATE: 2003-10-14
 25 <150> PRIOR APPLICATION NUMBER: JP2003-353491
 26 <151> PRIOR FILING DATE: 2003-10-14
 28 <150> PRIOR APPLICATION NUMBER: JP2004-18128
 W--> 29 <151> PRIOR FILING DATE: 2004-1-27
 31 <150> PRIOR APPLICATION NUMBER: JP2004-194088 *
 W--> 32 <151> PRIOR FILING DATE: 2004-6-30
 34 <160> NUMBER OF SEQ ID NOS: 37
 36 <170> SOFTWARE: PatentIn version 3.1
 38 <210> SEQ ID NO: 1
 39 <211> LENGTH: 1041
 40 <212> TYPE: DNA
 41 <213> ORGANISM: Escherichia coli
 43 <220> FEATURE:
 44 <221> NAME/KEY: CDS
 45 <222> LOCATION: (1)..(1041)
 46 <223> OTHER INFORMATION:
 W--> 48 <400> 1
 49 atg agc gac aac atc cgt gtt ggg ttg att ggg tat ggt tat gcg agc 48
 50 Met Ser Asp Asn Ile Arg Val Gly Leu Ile Gly Tyr Gly Tyr Ala Ser
 51 1 5 10 15
 52 aaa acc ttc cat gcg ccc ctg att gcg ggc acg ccc ggg cag gaa ctg 96
 53 Lys Thr Phe His Ala Pro Leu Ile Ala Gly Thr Pro Gly Gln Glu Leu
 54 20 25 30
 55 gcg gta atc tcc agc agt gat gaa aca aaa gta aaa gcc gac tgg cca 144
 56 Ala Val Ile Ser Ser Ser Asp Glu Thr Lys Val Lys Ala Asp Trp Pro
 57 35 40 45
 58 acg gtt acg gtt gtc tct gag ccg aag cat ctg ttt aac gat ccc aac 192
 59 Thr Val Thr Val Val Ser Glu Pro Lys His Leu Phe Asn Asp Pro Asn

RAW SEQUENCE LISTING

DATE: 04/27/2006

PATENT APPLICATION: US/10/576,030

TIME: 07:23:05

Input Set : A:\toyal49.001apc.txt

Output Set: N:\CRF4\04272006\J576030.raw

60	50	55	60	
61	ata gac ctg att gtc att cct aca ccc aac gat acc cat ttc ccg tta	240		
62	Ile Asp Leu Ile Val Ile Pro Thr Pro Asn Asp Thr His Phe Pro Leu			
63	65 70 75 80			
64	gcc aaa gcg gcg ctt gag gcg ggt aaa cat gtg gtc gtt gat aaa ccc	288		
65	Ala Lys Ala Ala Leu Glu Ala Gly Lys His Val Val Val Asp Lys Pro			
66	85 90 95			
67	ttt acc gtg aca ctg tca caa gcg cga gag ctg gat gcg ctg gca aaa	336		
68	Phe Thr Val Thr Leu Ser Gln Ala Arg Glu Leu Asp Ala Leu Ala Lys			
69	100 105 110			
70	agc ctg ggg cgt gtg ctg tct gta ttc cat aac cgt cgc tgg gat agc	384		
71	Ser Leu Gly Arg Val Leu Ser Val Phe His Asn Arg Arg Trp Asp Ser			
72	115 120 125			
73	gat ttc ttg acg cta aaa ggt tta ctc gcg gaa ggc gtg ctg ggt gaa	432		
74	Asp Phe Leu Thr Leu Lys Gly Leu Leu Ala Glu Gly Val Leu Gly Glu			
75	130 135 140			
76	gtt gct tac ttt gag tct cat ttt gac cgc ttc cgt ccg cag gtg cgc	480		
77	Val Ala Tyr Phe Glu Ser His Phe Asp Arg Phe Arg Pro Gln Val Arg			
78	145 150 155 160			
79	gat cgt tgg cgt gaa cag ggc ggc cca ggc agc ggt atc tgg tac gat	528		
80	Asp Arg Trp Arg Glu Gln Gly Gly Pro Gly Ser Gly Ile Trp Tyr Asp			
81	165 170 175			
82	tta gca cca cat ctt ctt gat cag gcc att acg ctt ttt ggt tta ccg	576		
83	Leu Ala Pro His Leu Leu Asp Gln Ala Ile Thr Leu Phe Gly Leu Pro			
84	180 185 190			
85	gtc agc atg acg gta gat ttg gca cag tta cgg ccc gga gcg cag tcg	624		
86	Val Ser Met Thr Val Asp Leu Ala Gln Leu Arg Pro Gly Ala Gln Ser			
87	195 200 205			
88	acc gat tat ttc cac gcc atc ttg tcc tat cca cag cgg cga gtc att	672		
89	Thr Asp Tyr Phe His Ala Ile Leu Ser Tyr Pro Gln Arg Arg Val Ile			
90	210 215 220			
91	tta cac ggt acc atg ctg gca gcc gct gag tca gca cgg tat atc gtg	720		
92	Leu His Gly Thr Met Leu Ala Ala Ala Glu Ser Ala Arg Tyr Ile Val			
93	225 230 235 240			
94	cat gga tcc cga ggc agt tat gtg aaa tat ggc ctc gat cca cag gaa	768		
95	His Gly Ser Arg Gly Ser Tyr Val Lys Tyr Gly Leu Asp Pro Gln Glu			
96	245 250 255			
97	gaa cgt ctg aaa aat ggc gag cgt cta ccg cag gaa gac tgg ggc tac	816		
98	Glu Arg Leu Lys Asn Gly Glu Arg Leu Pro Gln Glu Asp Trp Gly Tyr			
99	260 265 270			
100	gat atg cgt gat ggc gta ctt acc cgc gtg gaa ggt gag gaa cgt gtc	864		
101	Asp Met Arg Asp Gly Val Leu Thr Arg Val Glu Gly Glu Glu Arg Val			
102	275 280 285			
103	gaa gaa acg ctg ttg acg gtg cct ggg aat tat ccg gct tac tat gcg	912		
104	Glu Glu Thr Leu Leu Thr Val Pro Gly Asn Tyr Pro Ala Tyr Tyr Ala			
105	290 295 300			
106	gct att cgt gat gcg tta aat ggc gat ggt gaa aat ccg gtt ccg gca	960		
107	Ala Ile Arg Asp Ala Leu Asn Gly Asp Gly Glu Asn Pro Val Pro Ala			
108	305 310 315 320			

RAW SEQUENCE LISTING

DATE: 04/27/2006

PATENT APPLICATION: US/10/576,030

TIME: 07:23:05

Input Set : A:\toyal49.001apc.txt

Output Set: N:\CRF4\04272006\J576030.raw

```

109 agc cag gca atc cag gta atg gag ttg att gag ctg ggc atc gaa tcc      1008
110 Ser Gln Ala Ile Gln Val Met Glu Leu Ile Glu Leu Gly Ile Glu Ser
111           325           330           335
112 gcc aaa cat cgc gcg act ttg tgc ctt gca tga      1041
113 Ala Lys His Arg Ala Thr Leu Cys Leu Ala
114           340           345
117 <210> SEQ ID NO: 2
118 <211> LENGTH: 346
119 <212> TYPE: PRT
120 <213> ORGANISM: Escherichia coli
122 <400> SEQUENCE: 2
123 Met Ser Asp Asn Ile Arg Val Gly Leu Ile Gly Tyr Gly Tyr Ala Ser
124 1           5           10           15
125 Lys Thr Phe His Ala Pro Leu Ile Ala Gly Thr Pro Gly Gln Glu Leu
126           20           25           30
127 Ala Val Ile Ser Ser Ser Asp Glu Thr Lys Val Lys Ala Asp Trp Pro
128           35           40           45
129 Thr Val Thr Val Val Ser Glu Pro Lys His Leu Phe Asn Asp Pro Asn
130           50           55           60
131 Ile Asp Leu Ile Val Ile Pro Thr Pro Asn Asp Thr His Phe Pro Leu
132 65           70           75           80
133 Ala Lys Ala Ala Leu Glu Ala Gly Lys His Val Val Val Asp Lys Pro
134           85           90           95
135 Phe Thr Val Thr Leu Ser Gln Ala Arg Glu Leu Asp Ala Leu Ala Lys
136           100          105          110
137 Ser Leu Gly Arg Val Leu Ser Val Phe His Asn Arg Arg Trp Asp Ser
138           115          120          125
139 Asp Phe Leu Thr Leu Lys Gly Leu Leu Ala Glu Gly Val Leu Gly Glu
140           130          135          140
141 Val Ala Tyr Phe Glu Ser His Phe Asp Arg Phe Arg Pro Gln Val Arg
142 145          150          155          160
143 Asp Arg Trp Arg Glu Gln Gly Gly Pro Gly Ser Gly Ile Trp Tyr Asp
144           165          170          175
145 Leu Ala Pro His Leu Leu Asp Gln Ala Ile Thr Leu Phe Gly Leu Pro
146           180          185          190
147 Val Ser Met Thr Val Asp Leu Ala Gln Leu Arg Pro Gly Ala Gln Ser
148           195          200          205
149 Thr Asp Tyr Phe His Ala Ile Leu Ser Tyr Pro Gln Arg Arg Val Ile
150           210          215          220
151 Leu His Gly Thr Met Leu Ala Ala Ala Glu Ser Ala Arg Tyr Ile Val
152 225          230          235          240
153 His Gly Ser Arg Gly Ser Tyr Val Lys Tyr Gly Leu Asp Pro Gln Glu
154           245          250          255
155 Glu Arg Leu Lys Asn Gly Glu Arg Leu Pro Gln Glu Asp Trp Gly Tyr
156           260          265          270
157 Asp Met Arg Asp Gly Val Leu Thr Arg Val Glu Gly Glu Glu Arg Val
158           275          280          285
159 Glu Glu Thr Leu Leu Thr Val Pro Gly Asn Tyr Pro Ala Tyr Tyr Ala
160           290          295          300

```

RAW SEQUENCE LISTING

DATE: 04/27/2006

PATENT APPLICATION: US/10/576,030

TIME: 07:23:05

Input Set : A:\toyal49.001apc.txt

Output Set: N:\CRF4\04272006\J576030.raw

161 Ala Ile Arg Asp Ala Leu Asn Gly Asp Gly Glu Asn Pro Val Pro Ala
 162 305 310 315 320
 163 Ser Gln Ala Ile Gln Val Met Glu Leu Ile Glu Leu Gly Ile Glu Ser
 164 325 330 335
 165 Ala Lys His Arg Ala Thr Leu Cys Leu Ala
 166 340 345

169 <210> SEQ ID NO: 3

170 <211> LENGTH: 1170

171 <212> TYPE: DNA

172 <213> ORGANISM: Agrobacterium tumefaciens

174 <220> FEATURE:

175 <221> NAME/KEY: CDS

176 <222> LOCATION: (1)..(1170)

177 <223> OTHER INFORMATION:

W--> 179 <400> 3

180 atg tcc tcc gct aca aag aaa ttc gat agt cgc cgc att cgt ctc ggt 48
 181 Met Ser Ser Ala Thr Lys Lys Phe Asp Ser Arg Arg Ile Arg Leu Gly
 182 1 5 10 15
 183 atg gtc ggc ggc ggt cag ggc gcc ttc att ggc ggc gtg cat cgc atc 96
 184 Met Val Gly Gly Gln Gly Ala Phe Ile Gly Ala Val His Arg Ile
 185 20 25 30
 186 gcg gcc cgg ctg gat gac cgt tac gag ctg gtg gcc gga gcg ctt tcc 144
 187 Ala Ala Arg Leu Asp Asp Arg Tyr Glu Leu Val Ala Gly Ala Leu Ser
 188 35 40 45
 189 tcc gat ccc gcg cgt gcc gcc gcc tcg gca aca ctg ctc ggc att gcg 192
 190 Ser Asp Pro Ala Arg Ala Ala Ala Ser Ala Thr Leu Leu Gly Ile Ala
 191 50 55 60
 192 ccg gag cgc tcc tat gcc tcg ttc gag gac atg gcg gcg act gag gcc 240
 193 Pro Glu Arg Ser Tyr Ala Ser Phe Glu Asp Met Ala Ala Thr Glu Ala
 194 65 70 75 80
 195 ggc cgg gag gat ggc atc gag gca gtc gcc atc gtc acc ccc aac cat 288
 196 Gly Arg Glu Asp Gly Ile Glu Ala Val Ala Ile Val Thr Pro Asn His
 197 85 90 95
 198 ctg cat ttt gcc ccg tcc aag gcc ttt ctc gaa gcc ggc atc cac gtc 336
 199 Leu His Phe Ala Pro Ser Lys Ala Phe Leu Glu Ala Gly Ile His Val
 200 100 105 110
 201 atc tgc gac aag ccg gtg acc gcg acg ctg gaa gaa gcg aag gca ctg 384
 202 Ile Cys Asp Lys Pro Val Thr Ala Thr Leu Glu Glu Ala Lys Ala Leu
 203 115 120 125
 204 gcc ggg atc gtc aga gcc tcg gat agc ctt ttc gtg ctg acg cat aac 432
 205 Ala Gly Ile Val Arg Ala Ser Asp Ser Leu Phe Val Leu Thr His Asn
 206 130 135 140
 207 tac acc ggt tac gcc atg ctg cgg cag atg cgc gag atg atc gct gaa 480
 208 Tyr Thr Gly Tyr Ala Met Leu Arg Gln Met Arg Glu Met Ile Ala Glu
 209 145 150 155 160
 210 ggc gcc att ggc aag ctg cgc cat gtc cag gcc gaa tat gcg cag gac 528
 211 Gly Ala Ile Gly Lys Leu Arg His Val Gln Ala Glu Tyr Ala Gln Asp
 212 165 170 175
 213 tgg ctg acc gaa gcg gtc gaa aaa acc ggc gca aaa ggt gcg gaa tgg 576

RAW SEQUENCE LISTING

DATE: 04/27/2006

PATENT APPLICATION: US/10/576,030

TIME: 07:23:05

Input Set : A:\toyal49.001apc.txt

Output Set: N:\CRF4\04272006\J576030.raw

```

214 Trp Leu Thr Glu Ala Val Glu Lys Thr Gly Ala Lys Gly Ala Glu Trp
215      180      185      190
216 cgc acc gac ccc agc cgc tcc ggt gcg ggc ggc gcc atc ggc gat atc      624
217 Arg Thr Asp Pro Ser Arg Ser Gly Ala Gly Gly Ala Ile Gly Asp Ile
218      195      200      205
219 ggc act cac gcc ttc aac gct gct gcc ttt gtg acg ggt gaa atc ccc      672
220 Gly Thr His Ala Phe Asn Ala Ala Ala Phe Val Thr Gly Glu Ile Pro
221      210      215      220
222 agc agt ctt tat gcg gat ctc acg tcg ttt gtg ccg ggc cgg cag ctg      720
223 Ser Ser Leu Tyr Ala Asp Leu Thr Ser Phe Val Pro Gly Arg Gln Leu
224 225      230      235      240
225 gat gac agc gcc aat att ctt ttg cgt tac gac agt ggc gcc aag ggc      768
226 Asp Asp Ser Ala Asn Ile Leu Leu Arg Tyr Asp Ser Gly Ala Lys Gly
227      245      250      255
228 atg ctc tgg gca agc cag atc gcg gtc ggc aat gaa aat gcg ctg tca      816
229 Met Leu Trp Ala Ser Gln Ile Ala Val Gly Asn Glu Asn Ala Leu Ser
230      260      265      270
231 ctc cgg gtc tat ggc gac aag ggc ggg ctt gaa tgg cac cac cgg gtg      864
232 Leu Arg Val Tyr Gly Asp Lys Gly Gly Leu Glu Trp His His Arg Val
233      275      280      285
234 ccg gac gag ctg tgg ttc acg ccc tat ggc gag ccg aag cgg ctg att      912
235 Pro Asp Glu Leu Trp Phe Thr Pro Tyr Gly Glu Pro Lys Arg Leu Ile
236      290      295      300
237 acc cgc aac ggt gcg ggc gcg ggt gcc gct gca aac cgt gtc agt cgt      960
238 Thr Arg Asn Gly Ala Gly Ala Gly Ala Ala Ala Asn Arg Val Ser Arg
239 305      310      315      320
240 gtg cca tcc ggg cac ccg gag gga tat ctc gag ggt ttt gcg acg att      1008
241 Val Pro Ser Gly His Pro Glu Gly Tyr Leu Glu Gly Phe Ala Thr Ile
242      325      330      335
243 tac cgc gaa gcc gca gac gca atc atc gca aag agg gag gga gaa aca      1056
244 Tyr Arg Glu Ala Ala Asp Ala Ile Ile Ala Lys Arg Glu Gly Glu Thr
245      340      345      350
246 gcc gcc ggg gag gtg att tac ccc ggc atg gag gac ggc ctt gcg ggt      1104
247 Ala Ala Gly Glu Val Ile Tyr Pro Gly Met Glu Asp Gly Leu Ala Gly
248      355      360      365
249 ctc gca ttc atc gat gcg gcc gtt cgc tcc agc cag acc tcg acc tgg      1152
250 Leu Ala Phe Ile Asp Ala Ala Val Arg Ser Ser Gln Thr Ser Thr Trp
251      370      375      380
252 gtc ggg atc gac atc tag      1170
253 Val Gly Ile Asp Ile
254 385
257 <210> SEQ ID NO: 4
258 <211> LENGTH: 389
259 <212> TYPE: PRT
260 <213> ORGANISM: Agrobacterium tumefaciens
262 <400> SEQUENCE: 4
263 Met Ser Ser Ala Thr Lys Lys Phe Asp Ser Arg Arg Ile Arg Leu Gly
264 1      5      10      15
265 Met Val Gly Gly Gly Gln Gly Ala Phe Ile Gly Ala Val His Arg Ile

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/576,030

DATE: 04/27/2006
TIME: 07:23:06

Input Set : A:\toyal49.00lapc.txt
Output Set: N:\CRF4\04272006\J576030.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:29; N Pos. 9,15,18,27

Seq#:30; N Pos. 15,18,24

Seq#:32; N Pos. 7,17

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:15,16,17,18,19,20,21,22,23,24,25,26,29,30,31,32,33,34,35,36,37

VERIFICATION SUMMARY

DATE: 04/27/2006

PATENT APPLICATION: US/10/576,030

TIME: 07:23:06

Input Set : A:\toyal49.001apc.txt

Output Set: N:\CRF4\04272006\J576030.raw

L:16 M:270 C: Current Application Number differs, Replaced Application Number
L:17 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:29 M:256 W: Invalid Numeric Header Field, Wrong Prior FILING DATE:YYYY-MM-DD
L:32 M:256 W: Invalid Numeric Header Field, Wrong Prior FILING DATE:YYYY-MM-DD
L:48 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:46
L:179 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:3,Line#:177
L:325 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:5,Line#:323
L:471 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:7,Line#:469
L:607 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:9,Line#:605
L:753 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:11,Line#:751
L:899 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:13,Line#:897
L:1176 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:27,Line#:1174
L:1332 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29 after pos.:0
L:1355 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30 after pos.:0
L:1384 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32 after pos.:0